

# **Biotic Prediction**

Building the Computational Technology Infrastructure  
for Public Health and Environmental Forecasting

## **Software Requirements Document**

BP-SRD-1.0

Task Agreement: GSFC-CT-1

July 15, 2002

## Contents

<b>1</b>	<b>Overview</b>	<b>3</b>
1.1	Introduction . . . . .	3
1.2	Referenced Documents . . . . .	3
1.3	Document Overview . . . . .	3
<b>2</b>	<b>System Overview</b>	<b>4</b>
2.1	System Concept . . . . .	4
2.2	System Environment . . . . .	4
2.3	System Architecture . . . . .	4
<b>3</b>	<b>Functional Requirements</b>	<b>6</b>
3.1	Ingest . . . . .	6
3.1.1	Validation . . . . .	6
3.1.2	Field Data . . . . .	6
3.1.3	Remote Sensing Data . . . . .	6
3.1.4	Data Acquisition . . . . .	6
3.1.5	Monitoring and Reporting . . . . .	6
3.2	Pre-processing . . . . .	7
3.2.1	Merge Data . . . . .	7
3.3	Archive . . . . .	7
3.3.1	Database . . . . .	7
3.3.2	Internal File Store . . . . .	7
3.3.3	External Files . . . . .	7
3.4	Modeling . . . . .	7
3.5	Post-processing . . . . .	8
3.5.1	Reprojecting data . . . . .	8
3.5.2	Data Overlay . . . . .	8
3.5.3	Metadata . . . . .	8
3.6	User Interface . . . . .	8
3.6.1	Profile Database . . . . .	8
3.6.2	Roles . . . . .	8
3.6.3	Graphical User Interface . . . . .	9
3.6.4	File Management . . . . .	9
<b>4</b>	<b>Performance Requirements</b>	<b>9</b>
4.1	CT Project Scaling Milestones . . . . .	9
<b>5</b>	<b>Documentation Requirements</b>	<b>9</b>
5.1	Software Requirements Document . . . . .	9
5.2	Concept Of Operations . . . . .	9
5.3	Software Design Document . . . . .	9
5.4	Test Plan / Procedures Document . . . . .	10
5.5	User's Guide . . . . .	10
5.6	Software Maintenance Manual . . . . .	10

<b>6</b>	<b>Requirements Verification Matrix</b>	<b>11</b>
<b>7</b>	<b>Build Plan</b>	<b>15</b>
<b>A</b>	<b>Glossary</b>	<b>19</b>

## List of Tables

1	Referenced Documents . . . . .	3
2	Verification Methods . . . . .	11
3	Requirements Verification Matrix . . . . .	11
4	Build Plan . . . . .	15

## List of Figures

1	ISFS Functional Architecture . . . . .	5
---	--	---

# 1 Overview

## 1.1 Introduction

This project will develop the high-performance, computational technology infrastructure needed to analyze the past, present, and future geospatial distributions of living components of Earth environments. This involves moving a suite of key predictive, geostatistical biological models into a scalable, cost-effective cluster computing framework; collecting and integrating diverse Earth observational datasets for input into these models; and deploying this functionality as a Web-based service. The resulting infrastructure will be used in the ecological analysis and prediction of exotic species invasions. This new capability will be deployed at the USGS Midcontinent Ecological Science Center and extended to other scientific communities through the USGS National Biological Information Infrastructure program.

## 1.2 Referenced Documents

**Table 1.** Referenced Documents

Document Title	Version	Date
Software Engineering / Development Plan	1.0	2002-04-08
Concept of Operations	1.0	2002-07-15

## 1.3 Document Overview

This document, the *Software Requirements Document*, enumerates the software requirements for the *Invasive Species Forecasting System* (ISFS).

Some of these requirements are planned for immediate implementation in the software, and some are intended to be “design goals” for now and may be implemented in later builds of the software. By enumerating all of the requirements here, the software is described in its eventual, evolved state.

Section 2 provides a brief description of the system, its interfaces, and its functional subsystems. Sections 3, 4 and 5 lay out the explicit list of Functional, Performance, and Documentation Requirements. Further revisions of this document will break down these requirements into more detail and add more specific derived requirements as needed.

Section 6 provides a verification matrix with the method that will be used to verify that a requirement has been adequately attained. This matrix will be used as the basis for the *Software Test Plan* to be delivered later.

The Build Plan Matrix in Section 7 will lay out the time frame for the implementation of each requirement.

## 2 System Overview

### 2.1 System Concept

The ISFS will provide an environment for the application of different statistical models to geolocated reference datasets.

Users will be able to:

- Apply pre-defined models against existing datasets
- Add new or augment existing datasets
- Add new or alter existing models.

A primary goal of the evolved system is to optimize resource intensive computation by distributing it across a Linux cluster. This will minimize the time it takes to run the models and provide a better user experience.

In order to provide a baseline for comparison, the canonical system will:

- Show proof of concept for the canonical example and the validity of the example
- Establish metrics that can gauge performance of the system as a whole
- Provide an established baseline configuration from which future development can be compared
- Identify specific targets where computational processes can be enhanced to show scalar improvements in performance
- Identify the algorithms and tools that are used for processing

### 2.2 System Environment

The *Invasive Species Forecasting System* will, for this stage of development be hosted on a Linux server at the NASA Goddard Space Flight Center, Building 28. It will use COTS tools including ENVI image processing, IDL, Fortran programs, and shell scripting languages. Datasets are already loaded and in the correct formats for processing and benchmarking the canonical example.

The establishment of a dedicated server will allow comparison of performance and product quality as different operating parameters are modified. Once the benchmarks for the baseline system have been recorded; the exact configuration of the dedicated server will be archived. Then, the server will be used to prototype different operational scenarios as documented in the BP-CONOP document.

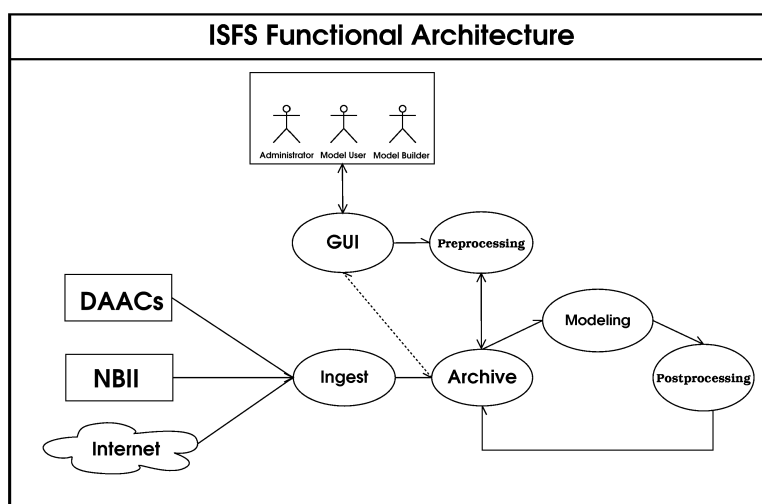
Subsequent mods to the baseline will be scrupulously documented along with any changes in performance that are measured or observed.

### 2.3 System Architecture

The ISFS conceptually consists of 6 functional elements:

1. Ingest
2. Pre-processing

3. Archive
4. Modeling
5. Post-processing
6. User Interface



**Figure 1.** ISFS Functional Architecture

## **3 Functional Requirements**

### **3.1 Ingest**

#### **3.1.1 Validation**

3.1.1.1 The system shall verify integrity, but not necessarily validate the quality of the data before ingest.

3.1.1.2 The system will attempt to ascertain that the data originated from an authoritative source.

#### **3.1.2 Field Data**

3.1.2.1 The system shall provide standard templates for ingesting field data in a tabular form.

3.1.2.2 The templates shall include all required fields to be captured.

3.1.2.3 The templates shall be in an accessible format (such as spreadsheet, database, or simple ASCII list).

#### **3.1.3 Remote Sensing Data**

3.1.3.1 The system shall support ingest from external satellite data archives (such as the Goddard DAAC).

3.1.3.2 The system shall support ingest of user-supplied satellite data or airborne imagery from digital files.

3.1.3.3 The system shall support ingest of user-supplied layers.

3.1.3.4 The system shall support ingest of user-supplied data files for ancillary layers.

#### **3.1.4 Data Acquisition**

3.1.4.1 The system shall support secured ftp-push for ingest of user supplied data.

3.1.4.2 The system shall support automated secured ftp pull from external archives.

3.1.4.3 The system shall provide accounting and logging by requiring users' name and password.

3.1.4.4 The system shall maintain a list of external archives and required data sets.

3.1.4.5 The interface between the system and each external archive will be thoroughly documented.

#### **3.1.5 Monitoring and Reporting**

3.1.5.1 The ingest subsystem shall monitor the number and volume of data brought into the system.

3.1.5.2 The ingest subsystem shall produce data reports broken down by location, user, and external archive.

3.1.5.3 The system shall generate and display associated metadata describing output files, runtime parameters, and performance statistics.

## **3.2 Pre-processing**

### **3.2.1 Merge Data**

3.2.1.1 The system shall merge ingested datasets.

3.2.1.2 The system shall perform resampling if the input data are not at the same resolution.

3.2.1.3 The data shall be converted to a common analysis format. [TBD]

3.2.1.4 The system shall merge ingested datasets into a data product that can be analyzed by the modeling subsystem.

## **3.3 Archive**

### **3.3.1 Database**

3.3.1.1 The Archive shall have a database that will store pointers to the archived files.

3.3.1.2 The database shall be able to refer to “internal” (stored in the local **File Store**) and “external” files.

3.3.1.3 All files (“internal” and “external”) shall be indexed with a unique file ID.

### **3.3.2 Internal File Store**

3.3.2.1 Files shall be stored in a logically arranged directory structure.

### **3.3.3 External Files**

3.3.3.1 For external files, the archive system shall store a pointer or URL that can be used to retrieve and stage the files for subsequent processing.

## **3.4 Modeling**

3.4.1 The system shall allow the ability to specify response and explanatory variables from the available databases.

3.4.2 The system shall provide graphical techniques (selection by rectangular or circular region) to explore the relationships between these variables.

3.4.3 The system shall have the ability to “fit” models through least squares (or other) optimization routines.

3.4.4 The system shall provide screening techniques to quantitatively assess which explanatory variables are related to the response variable (such as stepwise regression).

3.4.5 The system shall calculate geospatial statistics (such as variograms).

3.4.6 The system shall incorporate spatial structure into the modeling (such as generalized least squares or kriging).

3.4.7 The system shall be able to output model results and relevant model diagnostics.



3.4.8 The system shall allow the model builder to create new models.

3.4.9 The system shall allow the model user to select from an assortment of modeling techniques and to modify model parameters.

### **3.5 Post-processing**

#### **3.5.1 Reprojecting data**

3.5.1.1 The system shall display an image of the output data.

3.5.1.2 The system shall produce a data file suitable for reprojection by external COTS utilities.

#### **3.5.2 Data Overlay**

3.5.2.1 The system shall overlay output data with other layers as requested.

#### **3.5.3 Metadata**

3.5.3.1 Output data shall be packaged with appropriate metadata.

3.5.3.2 Each output data set shall be assigned a unique identifier.

### **3.6 User Interface**

#### **3.6.1 Profile Database**

3.6.1.1 The system shall maintain a profile database of users with:

1. User's status
2. User's preferences
3. User's role

#### **3.6.2 Roles**

3.6.2.1 The system shall support various roles that control access to various capabilities of the system:

1. "Administrator" — provides complete access to the system,
2. "Model Builder" — allows tailorability of the Ingest Subsystem and integration and use of new analytical routines to the Pre-Processing, Modeling, and Post-Processing Subsystems, and
3. "Model User" — allows tailorability of the Ingest Subsystem and use of existing analytical routines.

### **3.6.3 Graphical User Interface**

3.6.3.1 The system shall include a Graphical User Interface (“GUI”) to support user interaction with the system.

3.6.3.2 The GUI shall dynamically construct personalized web pages based on the Profile Database and the User’s Role.

3.6.3.3 The GUI shall display predictive map and uncertainty map output.

### **3.6.4 File Management**

3.6.4.1 The user shall have the option of saving run results with annotations in personal repository.

## **4 Performance Requirements**

### **4.1 CT Project Scaling Milestones**

4.1.1 Improve implementation of *PlantDiversity* to deliver canonical products 25X faster than the baseline implementation.

4.1.2 Improve implementation of *PlantDiversity* to accomodate 10X more sample data at 25X the time required in the baseline implementation and 10X larger area at 2.5X the time required in the baseline implementation.

## **5 Documentation Requirements**

### **5.1 Software Requirements Document**

A Requirements Document shall be provided which includes sufficient detail to provide requirement traceability appropriate for satisfying each milestone.

### **5.2 Concept Of Operations**

A Concept Of Operations shall be provided that describes the high level operations of the system.

### **5.3 Software Design Document**

System Design Documentation shall be provided which includes:

5.3.1 Sufficient detail to describe the architectural features, module interfaces, APIs, class hierarchies and functionality corresponding to the negotiated milestone.

5.3.2 A description of the problem class for which the code was designed;

5.3.3 An enumeration of the equations solved and their boundary conditions;

5.3.4 A description of the numerical methods used in the code;

5.3.5 A description of the parallelization techniques used.

- 5.3.6 References to easily accessible published papers or documents are acceptable as long as the specific sections of those documents which address this particular code are pointed out.

## **5.4 Test Plan / Procedures Document**

A Test Plan shall be provided which includes a suite of test cases that demonstrate that specific requirements for the milestone have been successfully completed.

Each test case must:

- 5.4.1 Specify all input and output data required to validate the tested functionality.
- 5.4.2 Describe the software components and functionality the test case verifies and the specific requirements satisfied.
- 5.4.3 Specify the system configuration, libraries, executables and any other test environment requirements necessary to successfully perform the test case.

## **5.5 User's Guide**

A Users Guide shall be provided with sufficient detail that other members of the Science Team and the user community can (a) scale the problem size or number of processors used up or down, and (b) solve a slightly different problem within the range for which the code was designed (a problem that requires only different initial data and the resetting of existing include-file parameters and flags), i.e., the user is able to compile and run the code for any such problem and be able to understand the results. This information shall include:

- 5.5.1 Instructions as to how to modify the source code/include file parameter settings and/or input file data values to change the problem size or number of PEs used to other acceptable settings.
- 5.5.2 A detailed description of all code input variables meanings and valid ranges sufficient to allow a user to adjust the problem being solved.
- 5.5.3 A detailed description of all code output variables sufficient to allow a user to understand the output of the code (including instructions for plotting the output, using standard software, if appropriate).
- 5.5.4 A detailed description of each example tutorial used to satisfy the requirements specified in Section 1, Part 3 of the Software Submission Criteria

## **5.6 Software Maintenance Manual**

A Maintenance Manual or equivalent documentation shall be provided with sufficient detail to allow members of the development team to maintain the source code, implement new features and enhance the functionality of the software system.

## 6 Requirements Verification Matrix

**Table 2.** Verification Methods

Verification Method	Description
Demonstration	The operation of the system, or a part of the system that relies on observable functional operation not requiring the use of instrumentation, special test equipment, or subsequent analysis.
Test	The operation of the system, or a part of the system, using instrumentation or other special test equipment to collect data for later analysis.
Analysis	The processing of accumulated data obtained from other qualification methods. Examples are reduction, interpolation, or extrapolation of test results.
Inspection	The visual examination of system components, documentation, etc.

**Table 3.** Requirements Verification Matrix

ID	Requirement	Ver. Meth.
<b>3.</b>	<b>Functional Requirements</b>	
<b>3.1</b>	<b>Ingest</b>	
<b>3.1.1</b>	<b>Validation</b>	
<b>3.1.1.1</b>	The system shall verify integrity, but not necessarily validate the quality of the data before ingest.	Test
<b>3.1.1.2</b>	The system will attempt to ascertain that the data originated from an authoritative source.	Demo./Anal.
<b>3.1.2</b>	<b>Field Data</b>	
<b>3.1.2.1</b>	The system shall provide standard templates for ingesting field data in a tabular form.	Demo.
<b>3.1.2.2</b>	The templates shall include all required fields to be captured.	Demo.
<b>3.1.2.3</b>	The templates shall be in an accessible format.	Demo./Anal.
<b>3.1.3</b>	<b>Remote Sensing Data</b>	
<b>3.1.3.1</b>	The system shall support ingest from external satellite data archives.	
<b>3.1.3.2</b>	The system shall support ingest of user-supplied satellite data or airborne imagery from digital files.	Demo.
<b>3.1.3.3</b>	The system shall support ingest of user-supplied layers.	Demo.
<b>3.1.3.4</b>	The system shall support ingest of user-supplied data files for ancillary layers.	Demo.
<b>3.1.4</b>	<b>Data Acquisition</b>	
<b>3.1.4.1</b>	The system shall support secured ftp-push for ingest of user supplied data.	Demo.
<b>3.1.4.2</b>	The system shall support automated secured ftp pull from external archives.	Demo.

<b>ID</b>	<b>Requirement</b>	<b>Ver. Meth.</b>
<b>3.1.4.3</b>	The system shall provide accounting and logging by requiring users name and password.	Demo.
<b>3.1.4.4</b>	The system shall maintain a list of external archives and required data sets.	Demo.
<b>3.1.4.5</b>	The interface between the system and each external archive will be thoroughly documented.	Insp.
<b>3.1.5</b>	<b>Monitoring and Reporting</b>	
<b>3.1.5.1</b>	The ingest subsystem shall monitor the number and volume of data brought into the system.	Test
<b>3.1.5.2</b>	The ingest subsystem shall produce data reports broken down by location, user, and external archive.	Test
<b>3.1.5.3</b>	The system shall generate and display associated metadata describing output files, runtime parameters, and performance statistics.	Demo./Test
<b>3.2</b>	<b>Pre-processing</b>	
<b>3.2.1</b>	<b>Merge Data</b>	
<b>3.2.1.1</b>	The system shall merge ingested datasets.	Demo.
<b>3.2.1.2</b>	The system shall perform resampling if the input data are not at the same resolution.	Demo./Anal.
<b>3.2.1.3</b>	The data shall be converted to a common analysis format.	Demo.
<b>3.2.1.4</b>	The system shall merge ingested datasets into a data product that can be analyzed by the modeling subsystem.	Demo.
<b>3.3</b>	<b>Archive</b>	
<b>3.3.1</b>	<b>Database</b>	
<b>3.3.1.1</b>	The Archive shall have a database that will store pointers to the archived files.	Demo.
<b>3.3.1.2</b>	The database shall be able to refer to internal (stored in the local File Store) and external files.	Demo.
<b>3.3.1.3</b>	All files (internal and external) shall be indexed with a unique file ID.	Demo.
<b>3.3.2</b>	<b>Internal File Store</b>	
<b>3.3.2.1</b>	Files shall be stored in a logically arranged directory structure.	Insp.
<b>3.3.3</b>	<b>External Files</b>	
<b>3.3.3.1</b>	For external files, the archive system shall store a pointer that can be used to retrieve and stage the files for subsequent processing.	Demo.
<b>3.4</b>	<b>Modeling</b>	
<b>3.4.1</b>	The system shall allow the ability to specify response and explanatory variables from the available databases.	Demo.
<b>3.4.2</b>	The system shall provide graphical techniques to explore the relationships between these variables.	Demo.
<b>3.4.3</b>	The system shall have the ability to fit models through least squares (or other) optimization routines.	Demo./Test/Anal.
<b>3.4.4</b>	The system shall provide screening techniques to quantitatively assess which explanatory variables are related to the response variable.	Demo./Test/Anal.

<b>ID</b>	<b>Requirement</b>	<b>Ver. Meth.</b>
<b>3.4.5</b>	The system shall calculate geospatial statistics.	Demo./Test/Anal.
<b>3.4.6</b>	The system shall incorporate spatial structure into the modeling.	Demo./Test/Anal.
<b>3.4.7</b>	The system shall be able to output model results and relevant model diagnostics.	Dmo.
<b>3.4.8</b>	The system shall allow the model builder to create new models.	Demo.
<b>3.4.9</b>	The system shall allow the model user to select from an assortment of modeling techniques and to modify model parameters.	Demo.
<b>3.5</b>	<b>Post-processing</b>	
<b>3.5.1</b>	<b>Reprojecting data</b>	
<b>3.5.1.1</b>	The system shall display an image of the output data.	Demo.
<b>3.5.1.2</b>	The system shall produce a data file suitable for reprojection by external COTS utilities.	Demo.
<b>3.5.2</b>	<b>Data Overlay</b>	
<b>3.5.2.1</b>	The system shall overlay output data with other layers as requested.	Demo.
<b>3.5.3</b>	<b>Metadata</b>	
<b>3.5.3.1</b>	Output data shall be packaged with appropriate metadata.	Demo.
<b>3.5.3.2</b>	Each output data set shall be assigned a unique identifier.	Demo.
<b>3.6</b>	<b>User Interface</b>	
<b>3.6.1</b>	<b>Profile Database</b>	
<b>3.6.1.1</b>	The system shall maintain a profile database of users.	Demo.
<b>3.6.2</b>	<b>Roles</b>	
<b>3.6.2.1</b>	The system shall support various roles that control access to various capabilities of the system.	Demo.
<b>3.6.3</b>	<b>Graphical User Interface</b>	
<b>3.6.3.1</b>	The system shall include a Graphical User Interface (GUI) to support user interaction with the system.	Demo.
<b>3.6.3.2</b>	The GUI shall dynamically construct personalized web pages based on the Profile Database and the User s Role.	Demo.
<b>3.6.3.3</b>	The GUI shall display predictive map and uncertainty map output.	Demo.
<b>3.6.4</b>	<b>File Management</b>	
<b>3.6.4.1</b>	The user shall have the option of saving run results with annotations in personal repository.	Demo.
<b>4.</b>	<b>Performance Requirements</b>	
<b>4.1</b>	<b>CT Project Scaling Milestones</b>	
<b>4.1.1</b>	Deliver canonical products 25X faster than the baseline implementation	Test/Anal.
<b>4.1.2</b>	Accommodate 10X more sample data at 25X the time required in the baseline implementation and 10X larger area at 2.5X the time required in the baseline implementation.	Test/Anal.
<b>5.</b>	<b>Documentation Requirements</b>	
<b>5.1</b>	Software Requirements Document	Insp.
<b>5.2</b>	Concept Of Operations	Insp.

<b>ID</b>	<b>Requirement</b>	<b>Ver. Meth.</b>
<b>5.3</b>	Software Design Document	Insp.
<b>5.3.1</b>	Sufficient detail to describe the software.	Insp.
<b>5.3.2</b>	A description of the problem class for which the code was designed.	Insp.
<b>5.3.3</b>	An enumeration of the equations solved and their boundary conditions.	Insp.
<b>5.3.4</b>	A description of the numerical methods used in the code.	Insp.
<b>5.3.5</b>	A description of the parallelization techniques used.	Insp.
<b>5.3.6</b>	References to easily accessible published papers or documents.	Insp.
<b>5.4</b>	Test Plan / Procedures Document	Insp.
<b>5.4.1</b>	Specify all input and output data required to validate the tested functionality.	Insp.
<b>5.4.2</b>	Describe the software components and functionality the test case verifies.	Insp.
<b>5.4.3</b>	Specify the system configuration, etc. to perform the test case.	Insp.
<b>5.5</b>	User's Guide	Insp.
<b>5.5.1</b>	Instructions as to how to modify the source code.	Insp.
<b>5.5.2</b>	A detailed description of all code input variables.	Insp.
<b>5.5.3</b>	A detailed description of all code output variables.	Insp.
<b>5.5.4</b>	A detailed description of each example tutorial.	Insp.
<b>5.6</b>	Software Maintenance Manual	Insp.

## 7 Build Plan

- B1** Build 1, Milestone F  
**B2** Build 2, Milestone G  
**B3** Build 3, Milestone K  
**DG** Design Goals, future builds

**Table 4.** Build Plan

ID	Requirement	B1 (F)	B2 (G)	B3 (K)	DG
<b>3.</b>	<b>Functional Requirements</b>				
<b>3.1</b>	<b>Ingest</b>				
<b>3.1.1</b>	<b>Validation</b>				
<b>3.1.1.1</b>	The system shall verify integrity, but not necessarily validate the quality of the data before ingest.	X	X	X	
<b>3.1.1.2</b>	The system will attempt to ascertain that the data originated from an authoritative source.				X
<b>3.1.2</b>	<b>Field Data</b>				
<b>3.1.2.1</b>	The system shall provide standard templates for ingesting field data in a tabular form.	X	X	X	
<b>3.1.2.2</b>	The templates shall include all required fields to be captured.	X	X	X	
<b>3.1.2.3</b>	The templates shall be in an accessible format.	X	X	X	
<b>3.1.3</b>	<b>Remote Sensing Data</b>				
<b>3.1.3.1</b>	The system shall support ingest from external satellite data archives.				X
<b>3.1.3.2</b>	The system shall support ingest of user-supplied satellite data or airborne imagery from digital files.	X			
<b>3.1.3.3</b>	The system shall support ingest of user-supplied layers.	X	X	X	
<b>3.1.3.4</b>	The system shall support ingest of user-supplied data files for ancillary layers.				
<b>3.1.4</b>	<b>Data Acquisition</b>				
<b>3.1.4.1</b>	The system shall support secured ftp-push for ingest of user supplied data.		X	X	
<b>3.1.4.2</b>	The system shall support automated secured ftp pull from external archives.				X
<b>3.1.4.3</b>	The system shall provide accounting and logging by requiring users name and password.		X	X	
<b>3.1.4.4</b>	The system shall maintain a list of external archives and required data sets.				X
<b>3.1.4.5</b>	The interface between the system and each external archive will be thoroughly documented.				X



ID	Requirement	B1 (F)	B2 (G)	B3 (K)	DG
<b>3.1.5</b>	<b>Monitoring and Reporting</b>				
<b>3.1.5.1</b>	The ingest subsystem shall monitor the number and volume of data brought into the system.		X	X	
<b>3.1.5.2</b>	The ingest subsystem shall produce data reports broken down by location, user, and external archive.			X	
<b>3.1.5.3</b>	The system shall generate and display associated metadata describing output files, runtime parameters, and performance statistics.			X	
<b>3.2</b>	<b>Pre-processing</b>				
<b>3.2.1</b>	<b>Merge Data</b>				
<b>3.2.1.1</b>	The system shall merge ingested datasets.				X
<b>3.2.1.2</b>	The system shall perform resampling if the input data are not at the same resolution.				X
<b>3.2.1.3</b>	The data shall be converted to a common analysis format.		X	X	
<b>3.2.1.4</b>	The system shall merge ingested datasets into a data product that can be analyzed by the modeling subsystem.				X
<b>3.3</b>	<b>Archive</b>				
<b>3.3.1</b>	<b>Database</b>				
<b>3.3.1.1</b>	The Archive shall have a database that will store pointers to the archived files.		X	X	
<b>3.3.1.2</b>	The database shall be able to refer to internal (stored in the local File Store) and external files.		X	X	
<b>3.3.1.3</b>	All files (internal and external) shall be indexed with a unique file ID.		X	X	
<b>3.3.2</b>	<b>Internal File Store</b>				
<b>3.3.2.1</b>	Files shall be stored in a logically arranged directory structure.		X	X	
<b>3.3.3</b>	<b>External Files</b>				
<b>3.3.3.1</b>	For external files, the archive system shall store a pointer that can be used to retrieve and stage the files for subsequent processing.		X	X	
<b>3.4</b>	<b>Modeling</b>				
<b>3.4.1</b>	The system shall allow the ability to specify response and explanatory variables from the available databases.	X	X	X	
<b>3.4.2</b>	The system shall provide graphical techniques to explore the relationships between these variables.		X	X	
<b>3.4.3</b>	The system shall have the ability to fit models through least squares (or other) optimization routines.	X	X	X	

ID	Requirement	B1 (F)	B2 (G)	B3 (K)	DG
3.4.4	The system shall provide screening techniques to quantitatively assess which explanatory variables are related to the response variable.		X	X	
3.4.5	The system shall calculate geospatial statistics.	X	X	X	
3.4.6	The system shall incorporate spatial structure into the modeling.				X
3.4.7	The system shall be able to output model results and relevant model diagnostics.	X	X	X	
3.4.8	The system shall allow the model builder to create new models.		X	X	
3.4.9	The system shall allow the model user to select from an assortment of modeling techniques and to modify model parameters.		X	X	
3.5	<b>Post-processing</b>				
3.5.1	<b>Reprojecting data</b>				
3.5.1.1	The system shall display an image of the output data.		X	X	
3.5.1.2	The system shall produce a data file suitable for reprojection by external COTS utilities.	X	X	X	
3.5.2	<b>Data Overlay</b>				
3.5.2.1	The system shall overlay output data with other layers as requested.			X	
3.5.3	<b>Metadata</b>				
3.5.3.1	Output data shall be packaged with appropriate metadata.		X	X	
3.5.3.2	Each output data set shall be assigned a unique identifier.		X	X	
3.6	<b>User Interface</b>				
3.6.1	<b>Profile Database</b>				
3.6.1.1	The system shall maintain a profile database of users.		X	X	
3.6.2	<b>Roles</b>				
3.6.2.1	The system shall support various roles that control access to various capabilities of the system.		X	X	
3.6.3	<b>Graphical User Interface</b>				
3.6.3.1	The system shall include a Graphical User Interface (GUI) to support user interaction with the system.		X	X	
3.6.3.2	The GUI shall dynamically construct personalized web pages based on the Profile Database and the User's Role.				X
3.6.3.3	The GUI shall display predictive map and uncertainty map output.		X	X	
3.6.4	<b>File Management</b>				
3.6.4.1	The user shall have the option of saving run results with annotations in personal repository.				X

ID	Requirement	B1 (F)	B2 (G)	B3 (K)	DG
<b>4.</b>	<b>Performance Requirements</b>				
<b>4.1</b>	<b>CT Project Scaling Milestones</b>				
<b>4.1.1</b>	Deliver canonical products 25X faster than the baseline implementation	X			
<b>4.1.2</b>	Accommodate 10X more sample data at 25X the time required in the baseline implementation and 10X larger area at 2.5X the time required in the baseline implementation.		X		
<b>5.</b>	<b>Documentation Requirements</b>				
<b>5.1</b>	Software Requirements Document	X	X	X	
<b>5.2</b>	Concept Of Operations	X	X	X	
<b>5.3</b>	Software Design Document	X	X	X	
<b>5.3.1</b>	Sufficient detail to describe the software.	X	X	X	
<b>5.3.2</b>	A description of the problem class for which the code was designed.	X	X	X	
<b>5.3.3</b>	An enumeration of the equations solved and their boundary conditions.	X	X	X	
<b>5.3.4</b>	A description of the numerical methods used in the code.	X	X	X	
<b>5.3.5</b>	A description of the parallelization techniques used.	X	X	X	
<b>5.3.6</b>	References to easily accessible published papers or documents.	X	X	X	
<b>5.4</b>	Test Plan / Procedures Document	X	X	X	
<b>5.4.1</b>	Specify all input and output data required to validate the tested functionality.	X	X	X	
<b>5.4.2</b>	Describe the software components and functionality the test case verifies.	X	X	X	
<b>5.4.3</b>	Specify the system configuration, etc. to perform the test case.	X	X	X	
<b>5.5</b>	User's Guide		X	X	
<b>5.5.1</b>	Instructions as to how to modify the source code.		X	X	
<b>5.5.2</b>	A detailed description of all code input variables.		X	X	
<b>5.5.3</b>	A detailed description of all code output variables.		X	X	
<b>5.5.4</b>	A detailed description of each example tutorial.		X	X	
<b>5.6</b>	Software Maintenance Manual			X	

## **A Glossary**

**BP** Biotic Prediction project  
**CT** Computational Technologies project  
**CONOP** Concept of Operations  
**COTS** Commercial Off The Shelf  
**CSU** Colorado State University  
**ESTO** Earth Science Technology Office  
**GSFC** Goddard Space Flight Center  
**GUI** Graphical User Interface  
**ISFS** Invasive Species Forecasting System  
**NREL** Natural Resources Ecology Laboratory  
**SEP** Software Engineering / Development Plan  
**URL** Uniform Resource Locator